Certificate of Mailing

Date of Deposit: January 16, 2004 Label Number:EV247566195US

I hereby certify under 37 C.F.R. § 1.10 that this correspondence is being deposited with the United States Postal Service as **"Express Mail Post Office to Addressee"** with sufficient postage on the date indicated above and is addressed to Mail Stop Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Elvis De La Cruz
Printed name of person mailing correspondence

Signature of person mailing correspondence

APPLICATION

FOR

UNITED STATES LETTERS PATENT

APPLICANT : FREDERICK M. AUSUBEL, DANIELLE A. GARSIN, COSTI D. SIFRI, STEPHEN B. CALDERWOOD, and ELEFTHERIOS MYLONAKIS

TITLE : ENTEROCOCCAL VIRULENCE FACTORS

ENTEROCOCCAL VIRULENCE FACTORS

Cross-Reference to Related Applications

This application is a continuation of International Application No. PCT/US02/22979, filed July 18, 2002, which was published in English under PCT Article 21(2), which claims the benefit of U.S. Provisional Application 60/306,212, filed July 18, 2001.

Background of the Invention

The invention relates to virulence factors and methods for identifying drugs that inhibit pathogen infections.

The gram-positive pathogens in the genus *Enterococcus* are an increasingly problematic source of nosocomial infections, in part due to multi-drug resistance. *Enterococcus* can cause diseases such as bacteremia and endocarditis. These pathogens can also infect the urinary tract and skin wounds in immunocompromised individuals. Infection can be fatal if the bacteria cannot be neutralized.

Despite their increasing prevalence as infectious agents, little is known about how these bacteria cause disease. Only cytolysin and aggregation substance have been studied rigorously enough to be established as virulence factors in mammalian models of *Enterococcus faecalis* pathogenesis. Other virulence factors, such as certain proteases, are believed to contribute to pathogenesis, but have not been studied adequately in mammalian model systems.

One reason that little is known about Enterococcal virulence factors is that the model systems used to study these bacteria, the favorite being a rabbit model of endocarditis, are expensive and unwieldy. Using a mammalian model system to screen for these virulence factors would be virtually impossible. Accordingly, there exists a need for straightforward, inexpensive, and reliable methods to identify Enterococcal virulence factors. Also needed, are easy, accurate screening methods that would greatly simplify the drug discovery process aimed at identifying molecules that inhibit Enterococcal pathogenicity or promote host resistance to this pathogen.

Summary of the Invention

The invention provides a novel approach to identifying Enterococcal virulence factors and for identifying compounds for treating bacterial pathogenesis. The invention further provides novel Enterococcal virulence factors, which serve as targets for drug discovery.

In one aspect, the invention features a method for identifying an Enterococcal virulence factor. The method, in general, involves the steps of: (a) exposing a nematode to a mutagenized Enterococcal pathogen; (b) determining whether the Enterococcal mutant infects the nematode, a reduction of disease in the nematode relative to that caused by the non-mutagenized Enterococcal pathogen indicating a mutation in an Enterococcal virulence factor; and (c) using the mutation as a marker for identifying the Enterococcal virulence factor. In preferred embodiments, the Enterococcal pathogen is *Enterococcus faecalis* (e.g., *Enterococcus faecalis* strain V583) and the nematode is *Caenorhabditis elegans* (e.g., a wild-type or mutant worm). In other preferred embodiments, the method utilizes an Enterococcal/*C. elegans* killing assay.

In another aspect, the invention features a method of identifying a compound that inhibits pathogenicity of an Enterococcal pathogen. The method, in general, involves the steps of: (a) providing a nematode infected with an Enterococcal pathogen; (b) contacting the infected nematode with a test compound; and (c) determining whether the test compound inhibits the pathogenicity of the Enterococcal pathogen in the nematode. In preferred embodiments, the Enterococcal pathogen is *Enterococcus faecalis* (e.g., *Enterococcus faecalis* strain V583) and the nematode is *Caenorhabditis elegans* (e.g., a wild-type or mutant worm). Preferably, the test compound is provided in a compound library. In other preferred embodiments the test compound is a small organic compound; or is a peptide, peptidomimetic, or antibody or fragment thereof. In still other preferred embodiments, the inhibition of pathogenicity is measured by an Enterococcal/*C. elegans* killing assay.

In yet another aspect, the invention features an isolated nematode (e.g., Caenorhabditis elegans), that includes an isolated Enterococcal pathogen. In preferred embodiments, the Enterococcal pathogen is *Enterococcus faecalis*, *Enterococcus faecalis* strain V583, *Enterococcus faecium*, or is a mutated Enterococcal pathogen.

In addition, we have identified and characterized a number of nucleic acid molecules and polypeptides that are involved in conferring pathogenicity and virulence to a pathogen. This discovery therefore provides a basis for drug-screening assays aimed at evaluating and identifying "anti-virulence" agents which are capable of blocking pathogenicity and virulence of a pathogen, e.g., by selectively switching pathogen gene expression on or off, or which inactivate or inhibit the activity of a polypeptide which is involved in the pathogenicity of a microbe. Drugs that target these molecules are useful as such anti-virulence agents.

Accordingly, the invention further features an isolated nucleic acid molecule including a sequence substantially identical to any one of the nucleic acid sequences encoding photolyase (Figure 5A) (SEQ ID NO:2), ScrR (Figure 6A) (SEQ ID NO:5), OppA (Figure 7A) (SEQ ID NO:8), TcaA (Figure 8A) (SEQ ID NO:11), ScrB (Figure 9A) (SEQ ID NO:14), RecQ (Figure 10A) (SEQ ID NO:17), LysR family of positive regulators (Figure 11A) (SEQ ID NO:20), XAA-His dipeptidase (carnosinase) (Figure 12A) (SEQ ID NO:23), SacU (Figure 13A) (SEQ ID NO:26), Pail (Figure 14A) (SEQ ID NO:29), or ORFs 1-8 of the shikimate amino acid biosynthetic gene cluster (Figures 15A, 15B, 15D, 15F, 15H, 15J, 15L, 15N, and 15P) (SEQ ID NOS:32-39). Preferably, the isolated nucleic acid molecule includes any of the above-described sequences or a fragment thereof; and is derived from a pathogen (e.g., from a bacterial pathogen such as *Enterococcus*). Additionally, the invention includes a vector and a cell, each of which includes at least one of the isolated nucleic acid molecules of the invention; and a method of producing a recombinant polypeptide involving providing a cell transformed with a nucleic acid molecule of the invention positioned for expression in the cell, culturing the transformed cell under conditions for expressing the nucleic acid molecule, and isolating a recombinant polypeptide. The invention further features recombinant polypeptides produced by such expression of an isolated nucleic acid molecule of the invention, and substantially pure antibodies that specifically recognize and bind to such recombinant polypeptides.

In an another aspect, the invention features a substantially pure polypeptide including an amino acid sequence that is substantially identical to the amino acid sequence of photolyase (Figure 5B) (SEQ ID NO:3), ScrR (Figure 6B) (SEQ ID NO:6), OppA (Figure 7B) (SEQ ID NO:9), TcaA (Figure 8B) (SEQ ID NO:12), ScrB (Figure 9B) (SEQ ID NO:15), RecQ (Figure 10B) (SEQ ID NO:18), LysR family of positive

regulators (Figure 11B) (SEQ ID NO:21), XAA-His dipeptidase (carnosinase) (Figure 12B) (SEQ ID NO:24), SacU (Figure 13B) (SEQ ID NO:27), Pail (Figure 14B) (SEQ ID NO:30), or ORFs 1-8 of the shikimate amino acid biosynthetic gene cluster (Figures 15A, 15C, 15E, 15G, 15I, 15K, 15M, 15O, and 15Q) (SEQ ID NO:40-47). Preferably, the substantially pure polypeptide includes any of the above-described sequences of a fragment thereof; and is derived from a pathogen (e.g., from a bacterial pathogen such as *Enterococcus)*.

In yet another related aspect, the invention features a method for identifying a compound which is capable of decreasing the expression of a pathogenic virulence factor (e.g., at the transcriptional or post-transcriptional levels), involving (a) providing a pathogenic cell expressing any one of the isolated nucleic acid molecules of the invention; and (b) contacting the pathogenic cell with a candidate compound, a decrease in expression of the nucleic acid molecule following contact with the candidate compound identifying a compound which decreases the expression of a pathogenic virulence factor. In preferred embodiments, the pathogenic cell infects a mammal (e.g., a human).

In yet another related aspect, the invention features a method for identifying a compound which binds a polypeptide, involving (a) contacting a candidate compound with a substantially pure polypeptide including any one of the amino acid sequences of the invention under conditions that allow binding; and (b) detecting binding of the candidate compound to the polypeptide.

In addition, the invention features a method of treating a pathogenic infection in a mammal, involving (a) identifying a mammal having a pathogenic infection; and (b) administering to the mammal a therapeutically effective amount of a composition which inhibits the expression or activity of a polypeptide encoded by any one of the nucleic acid molecules of the invention. In preferred embodiments, the pathogen is *Enterococcus*.

In yet another aspect, the invention features a method of treating a pathogenic infection in a mammal, involving (a) identifying a mammal having a pathogenic infection; and (b) administering to the mammal a therapeutically effective amount of a composition which binds and inhibits a polypeptide encoded by any one of the amino acid sequences of the invention. In preferred embodiments, the pathogenic infection is caused by *Enterococcus*.

The invention further features compounds or combinations of compounds used to treat or prevent microbial infections.

By "virulence factor" is meant a cellular component (e.g., a protein such as a transcription factor or a molecule) without which a pathogen is incapable of causing disease or infection in a eukaryotic host organism (e.g., a nematode or mammal). Such components are involved in the adaptation of the bacteria to a host (e.g., a nematode host), establishment of a bacterial infection, maintenance of a bacterial infection, and generation of the damaging effects of the infection to the host organism. Further, the phrase includes components that act directly on host tissue, as well as components which regulate the activity or production of other pathogenesis factors.

By "infection" or "infected" is meant an invasion or colonization of a host animal (e.g., nematode) by pathogenic bacteria that is damaging to the host.

By "inhibits pathogenicity of an Enterococcal pathogen" is meant the ability of a test compound to decrease, suppress, attenuate, diminish, arrest, or stabilize the development or progression of an Enterococcal-mediated disease or infection in a eukaryotic host organism. Preferably, such inhibition decreases pathogenicity by at least 5%, more preferably by at least 25%, and most preferably by at least 50% or more, as compared to symptoms in the absence of the test compound in any appropriate pathogenicity assay (for example, those assays described herein). In one particular example, inhibition may be measured by monitoring pathogenic symptoms in a nematode infected with an Enterococcal pathogen exposed to a test compound or extract, a decrease in the level of pathogenic symptoms relative to the level of symptoms in the host organism not exposed to the compound indicating compound-mediated inhibition of the Enterococcal pathogen.

By "isolated nucleic acid molecule" is meant a nucleic acid (e.g., a DNA) that is free of the genes which, in the naturally-occurring genome of the organism from which the nucleic acid molecule of the invention is derived, flank the gene. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or that exists as a separate molecule (for example, a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. In addition, the term includes an RNA molecule which is transcribed

from a DNA molecule, as well as a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence.

By "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (for example, glycosylation or phosphorylation).

By a "substantially pure polypeptide" is meant a polypeptide of the invention that has been separated from components which naturally accompany it. Typically, the polypeptide is substantially pure when it is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, a polypeptide of the invention. A substantially pure polypeptide of the invention may be obtained, for example, by extraction from a natural source (for example, a pathogen); by expression of a recombinant nucleic acid encoding such a polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, for example, column chromatography, polyacrylamide gel electrophoresis, or by HPLC analysis.

By "substantially identical" is meant a polypeptide or nucleic acid molecule exhibiting at least 25% identity to a reference amino acid sequence (for example, any one of the amino acid sequences described herein) or nucleic acid sequence (for example, any one of the nucleic acid sequences described herein). Preferably, such a sequence is at least 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, or 75% identical at the amino acid or nucleic acid level to the sequence used for comparison, more preferably 80%, or 85% and most preferably 90% or even 95% to 99% identical at the amino acid level or nucleic acid to the sequence used for comparison.

Sequence identity is typically measured using sequence analysis software (for example, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705, BLAST, BESTFIT, GAP, or PILEUP/PRETTYBOX programs). Such software matches identical or similar sequences by assigning degrees of homology to various substitutions, deletions, and/or other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. In an exemplary approach to determining the

degree of identity, a BLAST program may be used, with a probability score between e^{-3} and e^{-100} indicating a closely related sequence.

By "transformed cell" is meant a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a DNA molecule encoding (as used herein) a polypeptide of the invention.

By "positioned for expression" is meant that the DNA molecule is positioned adjacent to a DNA sequence which directs transcription and translation of the sequence (i.e., facilitates the production of, for example, a recombinant polypeptide of the invention, or an RNA molecule).

By "purified antibody" is meant antibody which is at least 60%, by weight, free from proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably 90%, and most preferably at least 99%, by weight, antibody. A purified antibody of the invention may be obtained, for example, by affinity chromatography using a recombinantly-produced polypeptide of the invention and standard techniques.

By "specifically binds" is meant a compound or antibody which recognizes and binds a polypeptide of the invention but which does not substantially recognize and bind other molecules in a sample, for example, a biological sample, which naturally includes a polypeptide of the invention.

By "derived from" is meant isolated from or having the sequence of a naturally-occurring sequence (e.g., a cDNA, genomic DNA, synthetic, or combination thereof).

The present invention provides a number of advantages. For example, the invention facilitates the identification of novel targets and therapeutic approaches for preparing therapeutic agents active on Enterococcal virulence factors and genes. The invention also provides long awaited advantages over a wide variety of standard screening methods used for distinguishing and evaluating the efficacy of a compound against Enterococcal pathogens. In one particular example, the screening methods described herein allow for the simultaneous evaluation of host toxicity as well as anti-Enterococcal potency in a simple *in vivo* screen. Moreover, the methods of the invention allow one to evaluate the ability of a compound to inhibit Enterococcal pathogenesis, and, at the same time, to evaluate the ability of the compound to stimulate and strengthen a host's response to Enterococcal pathogenic attack.

Accordingly, the methods of the invention provide a straightforward means to identify compounds that are both safe for use in eukaryotic host organisms (i.e., compounds which do not adversely affect the normal development and physiology of the organism) and efficacious against Enterococcal pathogenic microbes. In addition, the methods of the invention provide a route for analyzing virtually any number of compounds for anti-Enterococcal pathogenic effect with high-volume throughput, high sensitivity, and low complexity. The methods are also relatively inexpensive to perform and enable the analysis of small quantities of active substances found in either purified or crude extract form. Furthermore, the methods disclosed herein provide a means for identifying anti-pathogenic compounds which have the capability of crossing eukaryotic cell membranes and which maintain therapeutic efficacy in an *in vivo* method of administration. In addition, the above-described methods of screening are suitable for both known and unknown compounds and compound libraries.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

Brief Description of the Drawings

Figure 1 shows *C. elegans* killing by Enterococcal strains, E001, E002, E003, E006, E007, and E009.

Figure 2 shows *C. elegans* killing by *Enterococcus* strains E002, E006, and V583.

Figure 3 shows *C. elegans* killing by *Enterococcus* strains V583, OG1, OG1 (pAD1), OG1 (pCF10), and E007.

Figure 4 shows *C. elegans* killing by *Enterococcus* strains FA2-2, FA2-2 (pAM714), FA2-2(pAM771), and V583 containing plasmids pAD1 or pAD1-cyl .

Figure 5A shows the nucleic acid sequence of the Enterococcal virulence factor photolyase including the upstream and downstream untranslated regions (SEQ ID NO:1). The photolyase open reading frame (ORF) nucleic acid sequence (SEQ ID NO:2) extends from the start codon (ATG) to the stop (TAA) codon. The start and stop codons are indicated by highlighting.

Figure 5B shows the amino acid sequence of the Enterococcal virulence factor photolyase (SEQ ID NO:3).

Figure 6A shows the nucleic acid sequence of the Enterococcal virulence factor ScrR including the upstream and downstream untranslated regions (SEQ ID NO:4). The ScrR ORF nucleic acid sequence (SEQ ID NO:5) extends from the start codon (ATG) to the stop (TAA) codon. The start and stop codons are indicated by highlighting.

Figure 6B shows the amino acid sequence of the Enterococcal virulence factor ScrR (SEQ ID NO:6).

Figure 7A shows the nucleic acid sequence of the Enterococcal virulence factor OppA including the upstream and downstream untranslated regions (SEQ ID NO:7). The OppA ORF nucleic acid sequence (SEQ ID NO:8) extends from the start codon (ATG) to the stop (TAA) codon. The start and stop codons are indicated by highlighting.

Figure 7B shows the amino acid sequence of the Enterococcal virulence factor OppA (SEQ ID NO:9).

Figure 8A shows the nucleic acid sequence of an Enterococcal virulence factor, which is predicted to be a membrane protein related to TcaA, including the upstream and downstream untranslated regions (SEQ ID NO:10). The TcaA ORF nucleic acid sequence (SEQ ID NO:11) extends from the start codon (ATG) to the stop (TAA) codon. The start and stop codons are indicated by highlighting.

Figure 8B shows the amino acid sequence of an Enterococcal virulence factor which is predicted to be a membrane protein related to TcaA (SEQ ID NO:12).

Figure 9A shows the nucleic acid sequence of the Enterococcal virulence factor ScrB including the upstream and downstream untranslated regions (SEQ ID NO:13). The ScrB ORF nucleic acid sequence (SEQ ID NO:14) extends from the start codon (ATG) to the stop (TAG) codon. The start and stop codons are indicated by highlighting.

Figure 9B shows the amino acid sequence of the Enterococcal virulence factor ScrB (SEQ ID NO:15).

Figure 10A shows the nucleic acid sequence of the Enterococcal virulence factor RecQ including the upstream and downstream untranslated regions (SEQ ID NO:16). The RecQ ORF nucleic acid sequence (SEQ ID NO:17) extends from the start codon (ATG) to the stop (TAA) codon. The start and stop codons are indicated by highlighting.

Figure 10B shows the amino acid sequenceof the Enterococcal virulence factor RecQ (SEQ ID NO:18).

Figure 11A shows the nucleic acid sequence of an Enterococcal virulence factor, which is a member of the LysR family of positive regulators, including the upstream and downstream untranslated regions (SEQ ID NO:19). The LysR ORF nucleic acid sequence (SEQ ID NO:20) extends from the start codon (ATG) to the stop (TAG) codon. The start and stop codons are indicated by highlighting.

Figure 11B shows the amino acid sequence (SEQ ID NO:21) of an Enterococcal virulence factor which is a member of the LysR family of positive regulators.

Figure 12A shows the nucleic acid sequence of the Enterococcal virulence factor XAA-His dipeptidase (carnosinase) including the upstream and downstream untranslated regions (SEQ ID NO:22). The dipeptidase ORF nucleic acid sequence (SEQ ID NO:23) extends from the start codon (ATG) to the stop (TAA) codon. The start and stop codons are indicated by highlighting.

Figure 12B shows the amino acid sequence (SEQ ID NO:24) of the Enterococcal virulence factor XAA-His dipeptidase (carnosinase).

Figure 13A shows the nucleic acid sequence of the Enterococcal virulence factor SacU including the upstream and downstream untranslated regions (SEQ ID NO:25). The SacU ORF nucleic acid sequence (SEQ ID NO:26) extends from the start codon (ATG) to the stop (TAA) codon. The start and stop codons are indicated by highlighting.

Figure 13B shows the amino acid sequence (SEQ ID NO:27) of the Enterococcal virulence factor SacU.

Figure 14A shows the nucleic acid sequence of the Enterococcal virulence factor Pail including the upstream and downstream untranslated regions (SEQ ID NO:28). The Pail ORF nucleic acid sequence (SEQ ID NO:29) extends from the start codon (ATG) to the stop (TAA) codon. The start and stop codons are indicated by highlighting.

Figure 14B shows the amino acid sequence (SEQ ID NO:30) of the Enterococcal virulence factor Pail.

Figure 15A shows the nucleic acid sequence (SEQ ID NO:31) of an Enterococcal amino acid biosynthetic gene cluster that includes ORF1-ORF8 and the intergenic regions. ORF1 is from nucleotide 162-1028 (SEQ ID NO:32); ORF2 is from nucleotide

1065-2090 (SEQ ID NO:33); ORF3 is from nucleotide 2113-3189 (SEQ ID NO:34); ORF4 is from nucleotide 3191-4357 (SEQ ID NO:35); ORF5 is from nucleotide 4391-5485 (SEQ ID NO:36); ORF6 is from nucleotide 5492-6778 (SEQ ID NO:37); ORF7 is from nucleotide 6795-7304 (SEQ ID NO:38); and ORF8 is from nucleotide 7312-8160 (SEQ ID NO:39).

Figure 15B shows the nucleic acid sequence of ORF1.

Figure 15C shows the amino acid sequence of ORF1 (SEQ ID NO:40).

Figure 15D shows the nucleic acid sequence of ORF2.

Figure 15E shows the amino acid sequence of ORF2 (SEQ ID NO:41).

Figure 15F shows the nucleic acid sequence of ORF3.

Figure 15G shows the amino acid sequence of ORF3 (SEQ ID NO:42).

Figure 15H shows the nucleic acid sequence of ORF4.

Figure 15I shows the amino acid sequence of ORF4 (SEQ ID NO:43).

Figure 15J shows the nucleic acid sequence of ORF5.

Figure 15K shows the amino acid sequence of ORF5 (SEQ ID NO:44).

Figure 15L shows the nucleic acid sequence of ORF6.

Figure 15M shows the amino acid sequence of ORF6 (SEQ ID NO:45).

Figure 15N shows the nucleic acid sequence of ORF7.

Figure 15O shows the amino acid sequence of ORF7 (SEQ ID NO:46).

Figure 15P shows the nucleic acid sequence of ORF8.

Figure 15Q shows the amino acid sequence of ORF8 (SEQ ID NO:47).

Below we describe experimental evidence demonstrating that *Enterococcus* causes disease in the nematode *C. elegans*, and that *C. elegans* feeding on lawns of *Enterococcus faecalis* die over the course of a few days as a result of a pathogenic process. Data is also presented demonstrating that at least one known *E. faecalis* virulence factor, cytolysin, required for maximum virulence in mammalian models, enhances the killing of *C. elegans*, validating the use of a *C. elegans* host as a model for mammalian pathogenesis. The Enterococcal/*C. elegans* killing assay described herein therefore provides a useful system for identifying novel Enterococcal virulence factors, as well as for identifying compounds that either inhibit Enterococcal pathogenicity, promote a host's resistance to the pathogen, or both. Exemplary virulence factors identified in these assays are also disclosed. The following experimental examples are intended to illustrate, not limit, the scope of the claimed invention.

C. elegans/Enterococcus Killing Assays

To monitor *Enterococcus*-mediated killing, *C. elegans* assays were carried out as follows. Brain heart infusion (BHI) agar medium (Difco) was autoclaved and poured into 35 mm tissue culture plates (Fisher). Appropriate antibiotics were added to the medium before pouring that prevented growth of *E. coli*, but allowed growth of the particular *Enterococcus* strains being tested. For strains E001, E002, E003, E006 and E009, 12.5 μg/ml tetracycline was used. For strain V583, 200 μg/ml gentamycin was used. For strain E007, 50 μg/ml ampicillin was used. For strains OG1 and OG1 (pAD1), 250 μg/ml of spectinomycin was used. For strain OG1 (pCF10), 250 μg/ml spectinomycin and 12.5 μg/ml tetracycline were used. For strains FA2-2, FA2-2 (pAM714), and FA2-2 (pAM771), 50 μg/ml gentamycin was used.

Bacterial lawns of *Enterococcus* were prepared as follows. On the tissue culture plate, 2 ml of BHI was inoculated with a single colony of the appropriate strain, grown at 37 °C for four to five hours, and 10 μl of the culture was plated on each plate. The plates were incubated at 37 °C overnight, and then brought to room temperature for two to five hours. Thirty *C. elegans*, at the L4 larval stage, were then placed on the lawn from a plate of OP50 *E. coli*. The plates were incubated at 25 °C, and the number of worms found dead compared to the total number of plated worms was counted at approximately twenty-four hour intervals. Each experimental condition in the following experiments was done in triplicate and repeated at least twice.

C. elegans Killing By Enterococcal Clinical Isolates

Six different strains of *Enterococcus* were obtained from the clinical microbiology laboratory at Massachusetts General Hospital (Boston, MA) and were designated E001, E002, E003, E006, E007, and E009. Standard clinical methods were used to identify strains E003 and E007 as *Enterococcus faecium*, and strains E001, E002, E006, and E009 as *Enterococcus faecalis*. In general, *Enterococcus faecalis* causes approximately 80-90% of the Enterococcal infections in humans, and *Enterococcus faecium* causes approximately 10-20%.

Using the above-described killing assay protocol, the percentage of *C. elegans* dead as a function of time feeding on each of the *Enterococcus* strains was determined. As shown in Figure 1, clinical isolates E002 and E006 were found to kill *C. elegans*

most quickly, with a LT_{50} of about 100 hours. E001 and E009 killed more slowly, with a LT_{50} of about 150 hours. E003 and E007 did not cause any significant killing of *C. elegans*. These data suggest that *Enterococcus faecalis*, but not *Enterococcus faecium*, can kill *C. elegans*. Also, the differences between different strains of *Enterococcus faecalis* suggested that there might be identifiable genetic differences that caused the observed range in killing efficiency.

C. elegans Killing By *Enterococcus faecalis* strain V583

Strain V583 is a vancomycin-resistant clinical isolate of *E. faecalis*. It was first described in the late 1980's when vancomycin resistance emerged as a problem among Enterococcal infections. The genome of *E. faecalis* strain V583 is currently being sequenced by TIGR (The Institute for Genomic Research). The sequence is publicly available at (http://ftp.tigr.org/tdb/mdb/mdb.html).

Figure 2 shows the killing of *C. elegans* by strains E002, E006, and V583. In comparison to E002 and E006, *E. faecalis* strain V583 was found to kill just as effectively with a LT_{50} of about 100 hours (Figure 2).

C. elegans Killing By Isogenic *Enterococcus faecalis* Expressing Different Known Virulence Factors

E. faecalis contains a number of naturally occurring conjugative plasmids. Two such plasmids, called pAD1 (Jett et al., *Clinical Microbiol. Rev.* 7:462-478, 1994) and pCF10 (Leonard et al., *Proc. Natl. Acad. Sci.* 93:260-264, 1996) are well studied. A strain that does not have a particular plasmid (recipient) releases a peptide signal called a pheromone which, in turn, causes a strain that does contain the plasmid (donor) to produce aggregation substance (AS) on its surface. AS binds to *Enterococcus* binding substance (EBS) on the recipient, forming a mating aggregate which allows a copy of the plasmid to be conjugatively transferred from the donor to the recipient.

In addition to its role in plasmid conjugation, AS is also thought to play a role in pathogenesis by helping *E. faecalis* bind to host tissues. For example, strains producing AS bound more tightly to pig renal tubular cells than strains that did not produce AS. The genes for producing AS are located on both pAD1 and pCF10.

In addition to AS, cytolysin (Cyl) is another virulence factor that is capable of lysing both eukaryotic and other prokaryotic cells. The cytolysin operon is present on

pAD1, but not pCF10. When both AS and Cyl are expressed in rabbit endocarditis models, there is a significant increase in mortality.

To examine the role of AS and Cyl in nematode infection, isogenic strains of *E. faecalis* that were plasmid-free (strain OG1), harbored a plasmid containing genes for both AS and Cyl (strain OG1 containing pAD1), or harbored a plasmid containing only the gene for AS (strain OG1 containing pCF10) were examined in the above-described killing assay. *E. faecalis* strains V583 and E007 were used as controls. Results of these experiments are shown in Figure 3.

Strain OG1 containing pAD1 was found to kill *C. elegans* significantly faster than strain OG1 or strain OG1 with pCF10. These results appear to indicate that cytolysin contributes to pathogenesis in *C. elegans*, but that aggregation substance has little or no effect.

C. elegans Killing By Isogenic *Enterococcus* Strains Containing Conjugative Plasmids: pAD1 vs. pAD1-cyl

To examine whether the virulence factor, cytolysin, is responsible for the faster killing of the *C. elegans*, the killing rates of isogenic *E. faecalis* strains containing different mutations in pAD1 (Ike et al., *J. Bacteriol.* 172:155-163, 1990) were examined. As shown in Figure 4, the strain containing wild-type pAD1 (FA2-2 containing pAM714) was found to kill *C. elegans* faster than the plasmid-free strain or the strain containing pAD1 with a deletion in the promoter of the cytolysin operon (FA2-2 containing pAM771). These data indicated that the virulence factor, cytolysin, caused faster killing of *C. elegans*.

To summarize, we have developed a new pathogen/host model system employing *E. faecalis* and *C. elegans*. We have shown that different strains of *E. faecalis* kill *C. elegans* at different rates, and that *E. faecium* does not cause significant mortality. The sequenced strain of *E. faecalis* kills very effectively making it an ideal choice for mutagenesis studies. The known mammalian virulence factor, cytolysin, was also found to increase the rate of killing, suggesting that *C. elegans* is a valid model host for studying mammalian pathogenesis by *E. faecalis*. This model system provides a potentially valuable tool for identifying novel *E. faecalis* virulence factors, and for developing a better understanding of this problematic pathogen.

Nematode Screening Systems For Identifying Enteroccocal Virulence Factors

Based on the results described above showing that *E. faecalis* virulence factor is involved in pathogenicity of *C. elegans*, we have developed a method for identifying virulence determinants important for pathogenicity of *Enterococcus*. The screen, in general, utilizes the above-described Enterococcal/nematode killing assays and exploits the ability to readily screen thousands of randomly generated Enterococcal mutants. In addition to using wild-type host worms in the killing assays, mutants that are constipated or defecation defective, such as *aex-2* and *unc-25*, mutants that are grinding defective, such as *phm-2* and *eat-14*, and specific ABC transporter mutants such as *pgp-4* and *mrp-1* may be utilized as well.

In general, a strain of *Enterococcus* is mutated according to standard methods known in the art and then subsequently evaluated for its ability to induce disease in the nematode host organism. A mutagenized pathogen found to have diminished pathogenicity or which is rendered non-pathogenic is useful in the method of the invention. Such mutant pathogens are then used for identifying host-dependent or host-independent virulence factors responsible for pathogenicity according to methods known in the art.

The following is a working example of a virulence factor nematode screening system that utilizes the human clinical isolate *E. faecalis* strain V583 found to be infectious in the *C. elegans* nematode feeding model. Strain V583 is a vancomycin-resistant variety of *E. faecalis*, which contains plasmids amounting to an estimated 100 kb. It contains a set of seven genes spanning about 7 kb which contribute to its resistance to the vancomycin antiobiotic. The advantage of using a nematode as a host for studying this mammalian pathogen is the relative simplicity of identifying non-pathogenic *Enterococcus* mutants in the nematode.

In one preferred working example, in which survival is monitored, four to eight *C. elegans* worms (e.g., L4 larvae) are placed on a lawn of mutagenized *E. faecalis* strain V583, and survival is monitored after approximately one hundred to two hundred hours according to the methods described herein. An *Enterococcus* pathogen, such as *E. faecalis* strain V583, is mutated according to any standard procedure, e.g., standard *in vivo* or *in vitro* insertional/transponson mutagenesis methods (see, e.g., Ike et al., *J. Bacteriol.* 172:155-63, 1990; Munkenbeck et al., *Plasmid* 24:57-67, 1990; Kleckner et al., *J. Mol. Biol.* 116:125, 1977). Other methods are also available, e.g., chemical

mutagenesis, or directed mutagenesis of DNA. After approximately one hundred to one hundred fifty hours, very few or no live worms are found on a plate seeded with wild-type, pathogenic *E. faecalis* strain V583, whereas on a plate with mutagenized *E. faecalis* strain V583, increased survival (e.g., as determined by an increased LT_{50}) of the worms is observed. Thus, the ability of worms to grow in the presence of mutated *E. faecalis* strain V583 is an indication that a gene responsible for pathogenicity has been inactivated. The positions of the inactivating mutations are then identified using standard methods, (e.g., by polymerase chain reaction and sequencing of insertion/transposon junctions or by mapping), leading to the cloning and identification of the mutated virulence factor(s) (e.g., by nucleotide sequencing).

In another working example, in which survival and reproduction is monitored, two *C. elegans* worms (e.g., L4 hermaphrodite larvae) are placed on a lawn of mutagenized *E. faecalis* strain V583, and worm progeny is monitored. Strain V583 is mutated according to standard methods. After approximately one hundred to one hundred fifty hours, very few or no live worms are found on a plate seeded with wild-type, pathogenic *E. faecalis* strain V583, whereas on a plate with the V583 mutant, hundreds or thousands of live progeny of the initial two hermaphrodite worms are present. Thus, the ability of worms to grow and reproduce in the presence of mutated V583 is taken as an indication that a gene responsible for pathogenicity has been inactivated. The mutated virulence factor is then identified using standard methods.

Enterococcal Virulence Factors

Using the aforementioned screening assays, several novel Enterococcal virulence factors have been identified.

Nucleic acids encoding the Enterococcal virulence factors of the present invention were first identified using the following Enterococcal/nematode screening assay. To make a transposon library, OG1RF was transformed with pTV1-OK, which is a temperature sensitive vector containing the erythromycin-resistant transposon Tn917. A 50 ml culture from a single transformant was then grown in Brain-Heart Infusion (BHI) media at the nonselective temperature of 28°C to mid-log phase. The culture was then plated on BHI agar plates containing erythromycin and transposants selected for at the selective temperature of 42°C to select against the plasmid backbone and for chromosomal transposon insertions. The individual transposants were picked, grown in

individual wells of a 384-well plate and then frozen at -80°C to form the library. To screen how well these mutants killed *C. elegans* relative to wildtype OG1RF, transposon mutants were picked from the library, grown in liquid BHI media, and then plated as a lawn on 35 mm BHI-agar plates. Eight young adult nematodes were placed onto the *E. faecalis* lawn. The number of *C. elegans* alive at day two and day six were counted. Those mutants, causing significantly less death than average, were assayed again. In the second assay, each mutant was assayed in duplicate. The number of *C. elegans* alive was counted daily for six days. Mutants resulting in a significantly attenuated phenotype were assayed a third time in which each mutant was tested in triplicate using thirty worms. We sequenced the DNA flanking the transposon insertions of mutants that still displayed an attenuated phenotype. To do this we cultured mutants of interest in BHI media. The cultures were then boiled, solid material was removed by centrifugation, and the supernatant was used as the starting material for polymerase chain reaction (PCR) amplification of the genes of interest. For the PCR amplification, primers homologous to the transposon sequence and an arbitrary primer were utilized. The resulting material was next used in a second round of PCR with a second primer homologous to the transposon sequence and a primer homologous to the non-randomized sequence of the arbitrary primer. The resulting product was then sequenced and the DNA sequence flanking the transposon determined. The characterization of the identified virulence factors now follows.

Photolyase

A nucleic acid sequence (SEQ ID NO:2) (Figure 5A) encoding a photolyase of *Enterococcus faecalis* was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. First, the sequence information obtained by sequencing the region flanking the transposon insertion was used to query the *E. faecalis* unfinished genome sequence (available publicly at www.tigr.org). We found that the transposon had inserted within a large ORF. The ORF of interest was then blasted against the NCBI database of all publicly available sequences using the blastx program, which translates the queried sequence and then searches against the protein database. By this method we identified an ORF that was related to photolyase of *Bacillus firmus*. The *E. faecalis* photolyase

mutant was also tested using an *in vivo* mouse peritonitis assay and found to display an attenuated virulence phenotype.

In one embodiment, the invention encompasses a polypeptide that includes the photolyase amino acid sequence (SEQ ID NO:3) shown in Figure 5B. Enterococcal photolyase is 476 amino acids and shares significant amino acid identity to photolyases of other bacteria. In particular, *E. faecalis* photolyase and *Bacillus firmus* photolyase share 43% identity and 61% similarity. A fragment of SEQ ID NO:1 from about nucleotide 568 to nucleotide 1998 is useful as a hybridization probe. Photolyases are generally thought to be involved in the repair of light-damaged DNA.

The invention also encompasses Enterococcal photolyase variants. One preferred photolyase variant is one that has at least about 60%, more preferably at least 75%, and most preferably at least 90% amino acid sequence identity to the photolyase amino acid sequence.

ScrR

A nucleic acid sequence (SEQ ID NO:5) (Figure 6A) encoding ScrR of *Enterococcus faecalis* was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. The sequence analysis was done as described for photolyase (above). The ORF the transposon inserted into was found to have high similarity to scrR (the gene for the sucrose operon repressor) of various different bacterial species.

In one embodiment, the invention encompasses a polypeptide comprising the ScrR amino acid sequence (SEQ ID NO:6) shown in Figure 6B. Enterococcal ScrR is 325 amino acids in length and has very high identity to other ScrR sequences from other bacteria. For example, *E. faecalis* ScrR shares 52% similarity with Streptococcus mutants ScrR. A fragment of SEQ ID NO:4 from about nucleotide 662 to nucleotide 1629 is useful as a hybridization probe. ScrR functions to repress expression of the enzymes that degrade sucrose when no sucrose is available.

The invention also encompasses Enterococcal ScrR variants. One preferred ScrR variant is one that has at least about 60%, more preferably at least 75%, and most preferably at least 90% amino acid sequence identity to the ScrR amino acid sequence.

OppA

A nucleic acid sequence (SEQ ID NO:8) (Figure 7A) encoding an OppA-like protein of *Enterococcus faecalis* was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. The sequence analysis was done as described for photolyase (above). The sequence was found to have high similarity to OppA of *Listeria monocytogenes* and other bacterial peptide binding proteins.

In one embodiment, the invention encompasses a polypeptide that includes the OppA amino acid sequence (SEQ ID NO:9) shown in Figure 7B. Enterococcal OppA is 550 amino acids in length. The *E. faecalis* OppA protein shares 42% identity and 63% similarity to OppA of *Listeria monocytogenes*. A fragment of SEQ ID NO:7 from about nucleotide 421 to nucleotide 696 is useful as a hybridization probe. OppA of *Listeria monocytogenes* mediates the transport of oligopeptides and mutants in this gene are inhibited in intracellular survival and growth when infecting mice.

The invention also encompasses Enterococcal OppA variants. One preferred OppA variant is one that has at least about 60%, more preferably at least 75%, and most preferably at least 90% amino acid sequence identity to the OppA amino acid sequence.

TcaA

A nucleic acid sequence (SEQ ID NO:11) (Figure 8A) encoding a TcaA-like protein of *Enterococcus faecalis* was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. The sequence analysis was done as described for photolyase (above). The sequence was not strongly homologous to any proteins in the database. It had weak similarity to the TcaA protein of *Staphylococcus aureus* and several other known and hypothetical proteins that share the feature of being membrane-associated.

In one embodiment, the invention encompasses a polypeptide that includes the TcaA-like amino acid sequence (SEQ ID NO:12) shown in Figure 8B. Enterococcal TcaA-like protein is 443 amino acids in length and shares 20% identity and 36% similarity to *S. aureus* TcaA. A fragment of SEQ ID NO:10 from about nucleotide 1201 to nucleotide 2532 is useful as a hybridization probe. Mutations in *S. aureus* TcaA are associated with increased resistance to the antibiotic teicoplanin.

The invention also encompasses Enterococcal TcaA variants. One preferred TcaA variant is one that has at least about 60%, more preferably at least 75%, and most preferably at least 90% amino acid sequence identity to the TcaA amino acid sequence.

ScrB

A nucleic acid sequence (SEQ ID NO:14) (Figure 9A) encoding an ScrB of *Enterococcus faecalis* was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. The sequence analysis was done as described for photolyase (above). The protein sequence was strongly homologous to bacterial sucrose-6-phosphate hydrolases encoded by ScrB genes in other species. The *E. faecalis* ScrB mutant was also tested using an *in vivo* mouse peritonitis assay and found to display an attenuated virulence phenotype.

In one embodiment, the invention encompasses a polypeptidethat includes the ScrB amino acid sequence (SEQ ID NO:15) shown in Figure 9B. Enterococcal ScrB is 490 amino acids in length. It is highly homologous to other bacterial sucrases; for example it shares 48% identity and 64% similarity with ScrB of *Streptococcus sobrinus*. A fragment of SEQ ID NO:13 from about nucleotide 1018 to nucleotide 2487 is useful as a hybridization probe. Sucrose-6-phosphate hydrolases breakdown sucrose.

The invention also encompasses Enterococcal ScrB variants. One preferred ScrB variant is one that has at least about 60%, more preferably at least 75%, and most preferably at least 90% amino acid sequence identity to the ScrB amino acid sequence.

RecQ

A nucleic acid sequence (SEQ ID NO:17) (Figure 10A) encoding a RecQ of *Enterococcus faecalis* was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. The sequence analysis was done as described for photolyase (above). The protein sequence was most homologous to the ATP-dependent helicase RecQ of various bacterial species. The *E. faecalis* RecQ mutant was also tested using an *in vivo* mouse peritonitis assay and found to display an attenuated virulence phenotype.

In one embodiment, the invention encompasses a polypeptide that includes the RecQ amino acid sequence (SEQ ID NO:18) shown in Figure 10B. Enterococcal RecQ

is 448 amino acids in length. It was most homologous to RecQ of *Bacillus subtilis* with which it shares 37% identity and 56% similarity. RecQ is thought to unwind DNA during DNA repair processes. It has, for example, been shown to be involved in homologous recombination during pilin antigenic variation of the pathogen *Neisseria gonorrhoeae*. A fragment of SEQ ID NO:16 from about nucleotide 202 to nucleotide 1545 is useful as a hybridization probe.

The invention also encompasses Enterococcal RecQ variants. One preferred RecQ variant is one that has at least about 60%, more preferably at least 75%, and most preferably at least 90% amino acid sequence identity to the RecQ amino acid sequence.

LysR

A nucleic acid sequence (SEQ ID NO:20) (Figure 11A) encoding a positive regulator of the LysR family of *Enterococcus faecalis* was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. The sequence analysis was done as described for photolyase (above). The sequence was homologous to a family of positive regulators from various bacterial species which include CynR and LysR among others.

In one embodiment, the invention encompasses a polypeptide that includes the LysR-like amino acid sequence (SEQ ID NO:21) shown in Figure 11B. The Enterococcal LysR family member is 318 amino acids in length. The protein is most homologous to the CynR (a LysR-like regulator) from *Escherichia coli* with 27% identity and 48% similarity. These regulators are thought to positively regulate transcription of the operons they control under the appropriate conditions. A fragment of SEQ ID NO:19 from about nucleotide 847 to nucleotide 1800 is useful as a hybridization probe.

The invention also encompasses Enterococcal variants of this protein. One preferred variant is one that has at least about 60%, more preferably at least 75%, and most preferably at least 90% amino acid sequence identity to the LysR-like amino acid sequence.

Dipeptidase

A nucleic acid sequence (SEQ ID NO:23) (Figure 12A) encoding a dipeptidase of *Enterococcus faecalis* was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. The sequence analysis was done as described for photolyase (above). Specifically the dipeptidase appeared most closely related to bacterial carnosinases, which cleave aminoacyl-histidine dipeptides.

In one embodiment, the invention encompasses a polypeptide that includes the dipeptidase amino acid sequence (SEQ ID NO:24) shown in Figure 12B. This Enterococcal dipeptidase is 432 amino acids in length. The dipeptidase is most homologous to *Lactobacillus helveticus* carnosinase with 32% identity and 48% similarity. A fragment of SEQ ID NO:22 from about nucleotide 682 to nucleotide 1977 is useful as a hybridization probe.

The invention also encompasses Enterococcal dipeptidase variants. One preferred dipeptidase variant is one that has at least about 60%, more preferably at least 75%, and most preferably at least 90% amino acid sequence identity to the dipeptidase amino acid sequence.

SacU

A nucleic acid sequence (SEQ ID NO:26) (Figure 13A) encoding a SacU of *Enterococcus faecalis* was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. The sequence analysis was done as described for photolyase (above). The protein was most homologous to SacU, a two-component regulator of saccharolytic enzyme induction.

In one embodiment, the invention encompasses a polypeptide that includes the SacU amino acid sequence (SEQ ID NO:27) shown in Figure 13B. Enterococcal SacU is 282 amino acids in length. The protein is most homologous to *Bacillus subtilis* SacU with 37% identity and 57% similarity. A fragment of SEQ ID NO:25 from about nucleotide 1345 to nucleotide 2190 is useful as a hybridization probe.

The invention also encompasses Enterococcal SacU variants. One preferred SacU variant is one that has at least about 60%, more preferably at least 75%, and most preferably at least 90% amino acid sequence identity to the SacU amino acid sequence.

Pail

A nucleic acid sequence (SEQ ID NO:29) (Figure 14A) encoding Pail of *Enterococcus faecalis* was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. The sequence analysis was done as described for photolyase (above). The protein was most homologous to Pail, of various bacterial species, which is a negative transcriptional regulator. The *E. faecalis* Pail mutant was also tested using an in vivo mouse peritonitis assay and found to display an attenuated virulence phenotype.

In one embodiment, the invention encompasses a polypeptide that includes the Pail amino acid sequence (SEQ ID NO:30) shown in Figure 14B. Enterococcal Pail is 176 amino acids in length. The protein was most homologous to *Bacillus subtilis* Pail with 40% identity and 61% similarity. Pail affects sporulation and degradative-enzyme production in *Bacillus subtilis*. A fragment of SEQ ID NO:28 from about nucleotide 1225 to nucleotide 1749 is useful as a hybridization probe.

The invention also encompasses Enterococcal Pail variants. One preferred Pail variant is one that has at least about 60%, more preferably at least 75%, and most preferably at least 90% amino acid sequence identity to the Pail amino acid sequence.

Aromatic Amino Acid Biosynthetic Gene Cluster

A nucleic acid sequence (SEQ ID NO:31) encoding the aromatic amino acid biosynthetic gene cluster of *Enterococcus faecalis* (shown in Figure 15A) was identified from a mutant found to have decreased virulence in the above-described screening assay using a computer search for amino acid sequence alignments. First we blasted the sequence information obtained by sequencing the region flanking the transposon insertion against the *E. faecalis* unfinished genome sequence (available publicly at www.tigr.org). This search provided us with an exact match to our queried sequence and also gave us the DNA sequence upstream and downstream to this sequence. We then used this sequence to identify open reading frames (ORF's) existing within the sequence. We found that the transposon had inserted within the promoter region of a large operon containing eight ORFs. The operon of interest was then blasted against the NCBI database of all publicly available sequences using the blastx program, which translates the queried sequence and then searches against the protein database. This

search revealed that the *E. faecalis* aromatic biosynthesis gene cluster functions in virulence.

The first ORF nucleic acid sequence (SEQ ID NO:32) is aroE, shown in Figure 15B (from nucleotide 162-1028 of SEQ ID NO:11), which encodes the predicted amino acid sequence (SEQ ID NO:40) (shown in Figure 15C) of shikimate 5-dehydrogenase, a 288 amino acid polypeptide that catalyzes the fourth step in the biosynthesis of chorismate within the aromatic amino acid biosynthetic pathway.

The second ORF nucleic acid sequence (SEQ ID NO:33) is aroF, shown in Figure 15D (from nucleotide 1065-2090 of SEQ ID NO:11), which encodes the predicted amino acid sequence (SEQ ID NO:41) (Figure 15E) of phospho-2-dehydro-3-deoxyheptonate adolase, a 341 amino acid polypeptide that catalyzes the first step in the biosynthesis of chorismate within the aromatic amino acid biosynthetic pathway.

The third ORF nucleic acid sequence (SEQ ID NO:34) is aroB, shown in Figure 15F (from nucleotide 2113-3189 of SEQ ID NO:11), which encodes the predicted amino acid sequence (SEQ ID NO:42) (Figure 15G) of dehydroquinate synthase, a 358 amino acid polypeptide that catalyzes the second step in the biosynthesis of chorismate within the aromatic amino acid biosynthetic pathway.

The fourth ORF nucleic acid sequence (SEQ ID NO:35) is aroC, shown in Figure 15H (from nucleotide 3191-4357 of SEQ ID NO:11), which encodes the predicted amino acid sequence (SEQ ID NO:43) (Figure 15I) of chorismate synthase, a 388 amino acid polypeptide that catalyzes the seventh step in the biosynthesis of chorismate within the aromatic amino acid biosynthetic pathway.

The fifth ORF nucleic acid sequence (SEQ ID NO:36) is tyrA, shown in Figure 15J from nucleotide 4391-5485, which encodes the predicted amino acid sequence (SEQ ID NO:44) (Figure 15K) of chorismate mutase, a 364 amino acid polypeptide that is involved in tyrosine biosynthesis.

The sixth ORF nucleic acid sequence (SEQ ID NO:37), shown in Figure 15L, is aroA (from nucleotide 5492-6778 of SEQ ID NO:11), which encodes the predicted amino acid sequence (SEQ ID NO:45) (Figure 15M) of phosphoshikimate 1-carboxyvinyltransferase, a 428 amino acid polypeptide that catalyzes the sixth step in the biosynthesis of chorismate within the aromatic amino acid biosynthetic pathway.

The seventh ORF nucleic acid sequence (SEQ ID NO:38), shown in Figure 15N (from nucleotide 6795-7304 of SEQ ID NO:11) encodes the predicted amino acid

sequence (SEQ ID NO:46) shown in Figure 15O of shikimate kinase, a 169 amino acid polypeptide. This protein was identified as being related to the *Lactococcus lactis* shikimate kinase, having 40% identity and 63% similarity. Shikimate kinase, like the other proteins encoded by this operon, is involved in aminoacid biosynthesis.

The eighth ORF nucleic acid sequence (SEQ ID NO:39) shown in Figure 15P (from nucleotide 7312-8160 of SEQ ID NO:11) is pheA which encodes the predicted amino acid sequence (SEQ ID NO:47) of prephenate dehydratase, a 282 amino acid polypeptide involved in L-phenylalanine biosynthesis.

Isolation of Additional Virulence Genes

Based on the nucleotide and amino acid sequences described herein, the isolation of additional coding sequences of virulence factors from virtually any Enterococcal pathogen is made possible using standard strategies and techniques that are well known in the art. In addition, any pathogenic cell can serve as the nucleic acid source for the molecular cloning of any of the virulence genes described herein, and these sequences are identified as ones encoding a protein exhibiting pathogenicity-associated structures, properties, or activities. Examples of pathogenic bacteria include, without limitation, *Aerobacter*, *Aeromonas*, *Acinetobacter*, *Agrobacterium*, *Bacillus*, *Bacteroides*, *Bartonella*, *Bortella*, *Brucella*, *Calymmatobacterium*, *Campylobacter*, *Citrobacter*, *Clostridium*, *Cornyebacterium*, *Enterobacter*, *Escherichia*, *Francisella*, *Haemophilus*, *Hafnia*, *Helicobacter*, *Klebsiella*, *Legionella*, *Listeria*, *Morganella*, *Moraxella*, *Proteus*, *Providencia*, *Pseudomonas*, *Salmonella*, *Serratia*, *Shigella*, *Staphylococcus*, *Streptococcus*, *Treponema*, *Xanthomonas*, *Vibrio*, and *Yersinia*.

In one particular example of such an isolation technique, any one of the nucleotide sequences described herein may be used, together with conventional screening methods of nucleic acid hybridization screening. Such hybridization techniques and screening procedures are well known to those skilled in the art and are described, for example, in Benton and Davis (*Science* 196:180, 1977); Grunstein and Hogness (*Proc. Natl. Acad. Sci.*, USA 72:3961, 1975); Ausubel et al. (*Current Protocols in Molecular Biology*, Wiley Interscience, New York, 1997); Berger and Kimmel (supra); and Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York. In one particular example, all or part of the nucleic acid sequence of photolyase (Figure 5A), ScrR (Figure 6A), OppA (Figure 7A),

TcaA (Figure 8A), ScrB (Figure 9A), RecQ (Figure 10A), LysR family of positive regulators (Figure 11A), XAA-His dipeptidase (carnosinase)(Figure 12A), SacU (Figure 13A), Pail (Figure 14A), and ORFs 1-8 of the shikimate amino acid biosynthetic gene cluster (Figure 15A), may be used as a probe to screen a recombinant plant DNA library for genes having sequence identity to these genes. Hybridizing sequences are detected by plaque or colony hybridization according to standard methods.

Alternatively, using all or a portion of the amino acid sequence, one may readily design specific oligonucleotide probes, including degenerate oligonucleotide probes (i.e., a mixture of all possible coding sequences for a given amino acid sequence). These oligonucleotides may be based upon the sequence of either DNA strand and any appropriate portion of the sequence of the described protein(s). General methods for designing and preparing such probes are provided, for example, in Ausubel et al. (supra), and Berger and Kimmel, *Guide to Molecular Cloning Techniques*, 1987, Academic Press, New York. These oligonucleotides are useful for gene isolation, either through their use as probes capable of hybridizing to complementary sequences or as primers for various amplification techniques, for example, polymerase chain reaction (PCR) cloning strategies. If desired, a combination of different, detectably-labelled oligonucleotide probes may be used for the screening of a recombinant DNA library. Such libraries are prepared according to methods well known in the art, for example, as described in Ausubel et al. (supra), or they may be obtained from commercial sources.

As discussed above, sequence-specific oligonucleotides may also be used as primers in amplification cloning strategies, for example, using PCR. PCR methods are well known in the art and are described, for example, in PCR Technology, Erlich, ed., Stockton Press, London, 1989; *PCR Protocols: A Guide to Methods and Applications*, Innis et al., eds., Academic Press, Inc., New York, 1990; and Ausubel et al. (supra). Primers are optionally designed to allow cloning of the amplified product into a suitable vector, for example, by including appropriate restriction sites at the 5' and 3' ends of the amplified fragment (as described herein). If desired, nucleotide sequences may be isolated using the PCR "RACE" technique, or Rapid Amplification of cDNA Ends (see, e.g., Innis et al. (supra)). By this method, oligonucleotide primers based on a desired sequence are oriented in the 3' and 5' directions and are used to generate overlapping PCR fragments. These overlapping 3'- and 5'-end RACE products are combined to

produce an intact full-length cDNA. This method is described in Innis et al. (supra); and Frohman et al., *Proc. Natl. Acad. Sci.* USA 85:8998, 1988.

Partial virulence sequences, e.g., sequence tags, are also useful as hybridization probes for identifying full-length sequences, as well as for screening databases for identifying previously unidentified related virulence genes. Confirmation of a sequence's relatedness to a pathogenicity polypeptide may be accomplished by a variety of conventional methods including, but not limited to, functional complementation assays and sequence comparison of the gene and its expressed product. In addition, the activity of the gene product may be evaluated according to any of the techniques described herein, for example, the functional or immunological properties of its encoded product.

Once an appropriate sequence is identified, it is cloned according to standard methods and may be used, for example, for screening compounds that reduce the virulence of a pathogen.

Polypeptide Expression

In general, polypeptides of the invention may be produced by transformation of a suitable host cell with all or part of a polypeptide-encoding nucleic acid molecule or fragment thereof in a suitable expression vehicle.

Those skilled in the field of molecular biology will understand that any of a wide variety of expression systems may be used to provide the recombinant protein. The precise host cell used is not critical to the invention. A polypeptide of the invention may be produced in a prokaryotic host (e.g., *E. coli*) or in a eukaryotic host (e.g., *Saccharomyces cerevisiae*, insect cells, e.g., Sf21 cells, or mammalian cells, e.g., NIH 3T3, HeLa, or preferably COS cells). Such cells are available from a wide range of sources (e.g., the American Type Culture Collection, Rockland, MD; also, see, e.g., Ausubel et al., supra). The method of transformation or transfection and the choice of expression vehicle will depend on the host system selected. Transformation and transfection methods are described, e.g., in Ausubel et al. (supra); expression vehicles may be chosen from those provided, e.g., in *Cloning Vectors: A Laboratory Manual* (P.H. Pouwels et al., 1985, Supp. 1987).

One particular bacterial expression system for polypeptide production is the *E. coli* pET expression system (Novagen, Inc., Madison, WI). According to this expression

system, DNA encoding a polypeptide is inserted into a pET vector in an orientation designed to allow expression. Since the gene encoding such a polypeptide is under the control of the T7 regulatory signals, expression of the polypeptide is achieved by inducing the expression of T7 RNA polymerase in the host cell. This is typically achieved using host strains which express T7 RNA polymerase in response to IPTG induction. Once produced, recombinant polypeptide is then isolated according to standard methods known in the art, for example, those described herein.

Another bacterial expression system for polypeptide production is the pGEX expression system (Pharmacia). This system employs a GST gene fusion system which is designed for high-level expression of genes or gene fragments as fusion proteins with rapid purification and recovery of functional gene products. The protein of interest is fused to the carboxyl terminus of the glutathione S-transferase protein from *Schistosoma japonicum* and is readily purified from bacterial lysates by affinity chromatography using Glutathione Sepharose 4B. Fusion proteins can be recovered under mild conditions by elution with glutathione. Cleavage of the glutathione S-transferase domain from the fusion protein is facilitated by the presence of recognition sites for site-specific proteases upstream of this domain. For example, proteins expressed in pGEX-2T plasmids may be cleaved with thrombin; those expressed in pGEX-3X may be cleaved with factor Xa.

Once the recombinant polypeptide of the invention is expressed, it is isolated, e.g., using affinity chromatography. In one example, an antibody (e.g., produced as described herein) raised against a polypetide of the invention may be attached to a column and used to isolate the recombinant polypeptide. Lysis and fractionation of polypeptide-harboring cells prior to affinity chromatography may be performed by standard methods (see, e.g., Ausubel et al., supra).

Once isolated, the recombinant protein can, if desired, be further purified, e.g., by high performance liquid chromatography (see, e.g., Fisher, *Laboratory Techniques In Biochemistry And Molecular Biology*, eds., Work and Burdon, Elsevier, 1980).

Polypeptides of the invention, particularly short peptide fragments, can also be produced by chemical synthesis (e.g., by the methods described in *Solid Phase Peptide Synthesis*, 2nd ed., 1984 The Pierce Chemical Co., Rockford, IL).

These general techniques of polypeptide expression and purification can also be used to produce and isolate useful peptide fragments or analogs (described herein).

Antibodies

To generate antibodies, a coding sequence for a polypeptide of the invention may be expressed as a C-terminal fusion with glutathione S-transferase (GST) (Smith et al., Gene 67:31-40, 1988). The fusion protein is purified on glutathione-Sepharose beads, eluted with glutathione, cleaved with thrombin (at the engineered cleavage site), and purified to the degree necessary for immunization of rabbits. Primary immunizations are carried out with Freund's complete adjuvant and subsequent immunizations with Freund's incomplete adjuvant. Antibody titres are monitored by Western blot and immunoprecipitation analyses using the thrombin-cleaved protein fragment of the GST fusion protein. Immune sera are affinity purified using CNBr-Sepharose-coupled protein. Antiserum specificity is determined using a panel of unrelated GST proteins.

As an alternate or adjunct immunogen to GST fusion proteins, peptides corresponding to relatively unique immunogenic regions of a polypeptide of the invention may be generated and coupled to keyhole limpet hemocyanin (KLH) through an introduced C-terminal lysine. Antiserum to each of these peptides is similarly affinity purified on peptides conjugated to BSA, and specificity tested in ELISA and Western blots using peptide conjugates, and by Western blot and immunoprecipitation using the polypeptide expressed as a GST fusion protein.

Alternatively, monoclonal antibodies which specifically bind any one of the polypeptides of the invention are prepared according to standard hybridoma technology (see, e.g., Kohler et al., *Nature* 256:495, 1975; Kohler et al., *Eur. J. Immunol.* 6:511, 1976; Kohler et al., *Eur. J. Immunol.* 6:292, 1976; Hammerling et al., *In Monoclonal Antibodies and T Cell Hybridomas*, Elsevier, NY, 1981; Ausubel et al., supra). Once produced, monoclonal antibodies are also tested for specific recognition by Western blot or immunoprecipitation analysis (by the methods described in Ausubel et al., supra). Antibodies which specifically recognize the polypeptide of the invention are considered to be useful in the invention; such antibodies may be used, e.g., in an immunoassay. Alternatively monoclonal antibodies may be prepared using the polypeptide of the invention described above and a phage display library (Vaughan et al., *Nature Biotech.* 14:309-314, 1996).

Preferably, antibodies of the invention are produced using fragments of the polypeptide of the invention which lie outside generally conserved regions and appear likely to be antigenic, by criteria such as high frequency of charged residues. In one

specific example, such fragments are generated by standard techniques of PCR and cloned into the pGEX expression vector (Ausubel et al., supra). Fusion proteins are expressed in *E. coli* and purified using a glutathione agarose affinity matrix as described in Ausubel et al. (supra). To attempt to minimize the potential problems of low affinity or specificity of antisera, two or three such fusions are generated for each protein, and each fusion is injected into at least two rabbits. Antisera are raised by injections in a series, preferably including at least three booster injections.

Antibodies against any of the polypeptides described herein may be employed to treat bacterial infections.

Mouse Pathogenicity Screening Assays

To further evaluate the virulence of Enterococcal mutants identified in the above-described nematode screening assays, mouse pathogenicity/mortality studies are performed as follows. Female ICR Mice (Taconic, Germantown, NY or Charles River, Wilmington, MA) weighing 20 to 30 grams and housed 5 per cage, are used for evaluating the virulence of Enterococcal mutants. Mice, in groups of 6-10, are injected intraperitoneally with mutant bacteria in sterile rat fecal extracts (SRFE) as described below. The survival of mice receiving mutant bacteria is then compared to the survival of animals receiving an equal inoculum of wild-type bacteria (e.g., without a mutation). All animals have access to chow and water *ad libitum* throughout an experiment.

An exemplary bacterial inoculum is prepared as follows. *Enterococcus faecalis* OG1RF or Enterococcal mutants are grown overnight in BHI broth at 37 °C with gentle shaking. The cells are harvested by centrifugation, washed once with 0.9% saline, and then are resuspended in saline to an optical density of 2.2 to 2.8 at 600 nm. CFUs (colony-forming units) of cells suspensions are determined by plating serial dilutions onto BHI agar plates. Serial dilutions are prepared in saline and mixed with SRFE to the desired inoculum. For the preparation of SRFE, rat feces are dried, crushed, mixed with a volume of sterile distilled water three times that of the feces, and autoclaved. The resultant slurry is centrifuged, and the fecal extracts are removed aseptically. The extracts are then autoclaved and mixed with an enterococcal culture. Each inoculum is then diluted to a final 35% SRFE to yield the desired final inoculum.

Using a 25-gauge needle, mice are injected intraperitoneally with a 1 ml inoculum containing approximately 5×10^8 to 1×10^9 colony forming units of *E.*

faecalis or an Enterococcal mutant. After injection the animals are returned to their cages and monitored every eight hours for seven days. Surviving animals are then sacrificed and examined by autopsy. Control mice injected intraperitoneally with 1 ml of sterile SRFE are also examined.

Upon autopsy, bacteria are recovered from the kidneys or spleens under aseptic conditions. Peritoneal fluid and abdominal abscesses are also sampled for evaluation. Serial dilutions of the peritoneal fluid are prepared and 0.1 ml of each dilution is spread on agar plates for colony counts. Plates are then incubated under aerobic conditions for up to four days. BHI plates containing rifampin (for culturing Enterococcal wild-type *Enterococcus* OG1RF) or rifampin and erythromycin (for culturing Enterococcal mutants) are used for selection. Results are expressed, for example, by Kaplan-Meier curves and log rank test using STATA software (StataCorp. 1999. Stata Statistical Software: Release 6.0. College Station, TX: Stata Corporation).

Mutants showing a statistically significant difference or a statistical trend ($P \leq 0.20$) compared to the wild-type are, if desired, evaluated a second time. Mutants identified as having reduced virulence are taken as being useful in the invention.

Compound Screening Assays

As discussed above, our experimental results demonstrated that Enterococcal virulence factors are involved in pathogenicity of the nematode, *C. elegans*. Based on this discovery we have also developed a screening procedure for identifying therapeutic compounds (e.g., anti-pathogenicity pharmaceuticals) which can be used to inhibit the ability of the Enterococcal pathogen to cause infection. In general, the method involves screening any number of compounds for therapeutically-active agents by employing the Enterococcal/nematode killing system described herein. Based on our demonstration that these pathogens infect and kill *C. elegans*, it will be readily understood that a compound which interferes with the pathogenicity of *Enterococcus* in a nematode also provides an effective therapeutic agent in a mammal (e.g., a human patient). Whereas most antibiotics currently in medical use are either bactericidal or bacteriostatic, thus favoring resistant strains or mutants, the compounds identified in the screening procedures described herein do not kill the bacteria but instead render them non-pathogenic. Moreover, since the screening procedures of the invention are performed *in*

vivo, it is also unlikely that the identified compounds will be highly toxic to the host organism.

Accordingly, the methods of the invention simplify the evaluation, identification, and development of active agents such as drugs for the treatment of pathogenic diseases caused by Enterococcal microbes.

In addition, we have identified a number of enterococcal virulence factors that are involved in pathogenicity and that may therefore be used to screen for compounds that reduce the virulence of that organism, as well as other microbial pathogens. For example, the invention provides methods of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of a polypeptide or the gene expression of a nucleic acid sequence of the invention. The method of screening may involve high-throughput techniques.

Any number of methods are available for carrying out such screening assays. According to one approach, candidate compounds are added at varying concentrations to the culture medium of pathogenic cells expressing one of the nucleic acid sequences of the invention. Gene expression is then measured, for example, by standard Northern blot analysis (Ausubel et al., supra), using any appropriate fragment prepared from the nucleic acid molecule as a hybridization probe. The level of gene expression in the presence of the candidate compound is compared to the level measured in a control culture medium lacking the candidate molecule. A compound which promotes a decrease in the expression of the pathogenicity factor is considered useful in the invention; such a molecule may be used, for example, as a therapeutic to combat the pathogenicity of an infectious organism.

If desired, the effect of candidate compounds may, in the alternative, be measured at the level of polypeptide production using the same general approach and standard immunological techniques, such as Western blotting or immunoprecipitation with an antibody specific for a pathogenicity factor. For example, immunoassays may be used to detect or monitor the expression of at least one of the polypeptides of the invention in a pathogenic organism. Polyclonal or monoclonal antibodies (produced as described above) which are capable of binding to such a polypeptide may be used in any standard immunoassay format (e.g., ELISA, Western blot, or RIA assay) to measure the level of the pathogenicity polypeptide. A compound which promotes a decrease in the expression of the pathogenicity polypeptide is considered particularly useful. Again,

such a molecule may be used, for example, as a therapeutic to combat the pathogenicity of an infectious organism.

Alternatively, or in addition, candidate compounds may be screened for those which specifically bind to and inhibit a pathogenicity polypeptide of the invention. The efficacy of such a candidate compound is dependent upon its ability to interact with the pathogenicity polypeptide. Such an interaction can be readily assayed using any number of standard binding techniques and functional assays (e.g., those described in Ausubel et al., supra). For example, a candidate compound may be tested *in vitro* for interaction and binding with a polypeptide of the invention and its ability to modulate pathogenicity may be assayed by any standard assays (e.g., those described herein).

In one particular example, a candidate compound that binds to a pathogenicity polypeptide may be identified using a chromatography-based technique. For example, a recombinant polypeptide of the invention may be purified by standard techniques from cells engineered to express the polypeptide (e.g., those described above) and may be immobilized on a column. A solution of candidate compounds is then passed through the column, and a compound specific for the pathogenicity polypeptide is identified on the basis of its ability to bind to the pathogenicity polypeptide and be immobilized on the column. To isolate the compound, the column is washed to remove non-specifically bound molecules, and the compound of interest is then released from the column and collected. Compounds isolated by this method (or any other appropriate method) may, if desired, be further purified (e.g., by high performance liquid chromatography). In addition, these candidate compounds may be tested for their ability to render a pathogen less virulent (e.g., as described herein). Compounds isolated by this approach may also be used, for example, as therapeutics to treat or prevent the onset of a pathogenic infection, disease, or both. Compounds which are identified as binding to pathogenicity polypeptides with an affinity constant less than or equal to 10 mM are considered particularly useful in the invention.

Potential antagonists include organic molecules, peptides, peptide mimetics, polypeptides, and antibodies that bind to a nucleic acid sequence or polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also include small molecules that bind to and occupy the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Other potential antagonists include antisense molecules.

Each of the DNA sequences provided herein may also be used in the discovery and development of antipathogenic compounds (e.g., antibiotics). The encoded protein, upon expression, can be used as a target for the screening of antibacterial drugs. Additionally, the DNA sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The invention also provides the use of the polypeptide, polynucleotide, or inhibitor of the invention to interfere with the initial physical interaction between a pathogen and mammalian host responsible for infection. In particular the molecules of the invention may be used: in the prevention of adhesion and colonization of bacteria to mammalian extracellular matrix proteins; to extracellular matrix proteins in wounds; to block mammalian cell invasion; or to block the normal progression of pathogenesis.

The antagonists and agonists of the invention may be employed, for instance, to inhibit and treat a variety of bacterial infections.

Optionally, compounds identified in any of the above-described assays may be confirmed as useful in conferring protection against the development of a pathogenic infection in any standard animal model (e.g., the mouse peritonitis assay described herein) and, if successful, may be used as anti-pathogen therapeutics (e.g, antibiotics).

In general, the chemical screening methods of the invention provide a straightforward means for selecting natural product extracts or compounds of interest from a large population which are further evaluated and condensed to a few active and selective materials. Constituents of this pool are then purified and evaluated in the methods of the invention to determine their anti-pathogenic activity.

Test Extracts and Compounds

In general, novel anti-pathogenic drugs are identified from large libraries of both natural product or synthetic (or semi-synthetic) extracts or chemical libraries according to methods known in the art. The screening method of the present invention is appropriate and useful for testing compounds from a variety of sources for possible anti-pathogenic activity. The initial screens may be performed using a diverse library of compounds, but the method is suitable for a variety of other compounds and compound libraries. Such compound libraries can be combinatorial libraries, natural product

libraries, or other small molecule libraries. In addition, compounds from commercial sources can be tested, as well as commercially available analogs of identified inhibitors.

For example, those skilled in the field of drug discovery and development will understand that the precise source of test extracts or compounds is not critical to the screening procedure(s) of the invention. Accordingly, virtually any number of chemical extracts or compounds can be screened using the methods described herein. Examples of such extracts or compounds include, but are not limited to, plant-, fungal-, prokaryotic- or animal-based extracts, fermentation broths, and synthetic compounds, as well as modification of existing compounds. Numerous methods are also available for generating random or directed synthesis (e.g., semi-synthesis or total synthesis) of any number of chemical compounds, including, but not limited to, saccharide-, lipid-, peptide-, and nucleic acid-based compounds. Synthetic compound libraries are commercially available from Brandon Associates (Merrimack, NH) and Aldrich Chemical (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant, and animal extracts are commercially available from a number of sources, including Biotics (Sussex, UK), Xenova (Slough, UK), Harbor Branch Oceangraphics Institute (Ft. Pierce, FL), and PharmaMar, U.S.A. (Cambridge, MA). In addition, natural and synthetically produced libraries are produced, if desired, according to methods known in the art, e.g., by standard extraction and fractionation methods. Furthermore, if desired, any library or compound is readily modified using standard chemical, physical, or biochemical methods.

In addition, those skilled in the art of drug discovery and development readily understand that methods for dereplication (e.g., taxonomic dereplication, biological dereplication, and chemical dereplication, or any combination thereof) or the elimination of replicates or repeats of materials already known for their anti-pathogenic activity should be employed whenever possible.

When a crude extract is found to have anti-pathogenic activity, further fractionation of the positive lead extract is necessary to isolate chemical constituents responsible for the observed effect. Thus, the goal of the extraction, fractionation, and purification process is the careful characterization and identification of a chemical entity within the crude extract having anti-pathogenic activity. Methods of fractionation and purification of such heterogenous extracts are known in the art. If desired, compounds

shown to be useful agents for the treatment of pathogenicity are chemically modified according to methods known in the art.

Since many of the compounds in libraries such as combinatorial and natural products libraries, as well as in natural products preparations, are not characterized, the screening methods of this invention provide novel compounds which are active as inhibitors or inducers in the particular screens, in addition to identifying known compounds which are active in the screens. Therefore, this invention includes such novel compounds, as well as the use of both novel and known compounds in pharmaceutical compositions and methods of treating.

Exemplary High Throughput Screening Systems

To evaluate the efficacy of a molecule or compound in promoting host resistance to, or inhibiting pathogenicity of *Enterococcus*, a number of high throughput assays may be utilized.

For example, to enable mass screening of large quantities of natural products, extracts, or compounds in an efficient and systematic fashion, *Caenorhabditis elegans*, (e.g., L4 hermaphrodite larvae or a mutant worm such as *aex-2*, *unc-25*, *phm-2*, *eat-14*, *pgp-4*, or *mrp-1*), are cultured in wells of a microtiter plate, facilitating the semiautomation of manipulations and full automation of data collection. As is discussed above, *E. faecalis* infects and kills *C. elegans*. If *E. faecalis* has diminished pathogenicity, then L4 worms live, develop into adult hermaphrodites, and produce thousands of live progeny. Accordingly, if *C. elegans* is incubated with the pathogen, the worms will die, unless a compound is present to reduce *E. faecalis* pathogenicity. The presence of such live progeny is easily detected using a variety of methods, including visual screening with standard microscopes.

To evaluate the ability of a test compound or extract to promote a host's resistance to a pathogen or to repress pathogenicity of a pathogen, a test compound or extract is inoculated at an appropriate dosage into an appropriate agar medium (e.g., BHI or M17 (Difco)) seeded with an appropriate amount of an overnight culture of a pathogen, e.g., *E. faecalis*. If desired, various concentrations of the test compound or extract can be inoculated to assess dosage effect on both the host and the pathogen. Control wells are inoculated with non-pathogenic bacteria (negative control) or a pathogen in the absence of a test compound or extract (positive control). Plates are then

incubated twenty-four hours at 37 °C to facilitate the growth of the pathogen. Microtiter dishes are subsequently cooled to 25 °C, and two *C. elegans* L4 hermaphrodite larva are added to the plate and incubated at 25 °C, the upper limit for normal physiological integrity of *C. elegans*. At an appropriate time interval, e.g., one hundred to two hundred hours, wells are examined for surviving worms, the presence of progeny, or both, e.g., by visual screening or monitoring motion of worms using a motion detector.

In another working example, *Enterococcus*-mediated killing of *C. elegans* is carried out as follows. Brain heart infusion (BHI) agar medium (Difco) is autoclaved and poured into 35 mm tissue culture plates (Fisher). Appropriate antibiotics are added to the medium before pouring to prevent growth of *E. coli*, but allow for the growth of the particular *Enterococcus* strains being tested. A test compound or compound library is also added to the medium. On the tissue culture plate, 2 ml of BHI is inoculated with a single colony of the appropriate strain, grown at 37 °C for four to five hours, and 10 µl of the culture is plated on each plate. The plates are incubated at 37 °C overnight, and then brought to room temperature for two to five hours. Thirty *C. elegans*, at the L4 larval stage, are then placed on the lawn from a plate of OP50 *E. coli*. The plates are incubated at 25 °C, and the number of worms found dead compared to the total number of plated worms are then counted at approximately twenty-four hour intervals. Each experimental condition is done in triplicate and repeated at least twice. At an appropriate time interval plates are examined for surviving worms.

Comparative studies between treated and control worms (or larvae) are used to determine the relative efficacy of the test molecule or compound in promoting the host's resistance to the pathogen or inhibiting the virulence of the pathogen. A test compound which effectively stimulates, boosts, enhances, increases, or promotes the host's resistance to the pathogen or which inhibits, inactivates, suppresses, represses, or controls pathogenicity of the pathogen, and does not significantly adversely affect the normal physiology, reproduction, or development of the worms is considered useful in the invention.

Vaccines

The invention also includes vaccine compositions or formulations including the polypeptides or nucleic acid sequences of the invention. For example, the polypeptides of the invention may be used as an antigen for vaccination of a host to produce specific

antibodies which protect against invasion or infection or both of bacteria. The invention therefore includes a vaccine formulation which includes an immunogenic recombinant polypeptide of the invention together with a suitable carrier. For example, the invention also includes a vaccine formulation which includes an immunogenic recombinant protein of the invention together with a suitable carrier.

Use

The methods of the invention provide a simple means for identifying Enterococcal virulence factors and compounds capable of either inhibiting pathogenicity or enhancing an organism's resistance capabilities to such pathogens. Accordingly, a chemical entity discovered to have medicinal value using the methods described herein are useful as either drugs, or as information for structural modification of existing anti-pathogenic compounds, e.g., by rational drug design.

For therapeutic uses, the compositions or agents identified using the methods disclosed herein may be administered systemically, for example, formulated in a pharmaceutically-acceptable buffer such as physiological saline. Preferable routes of administration include, for example, subcutaneous, intravenous, interperitoneally, intramuscular, or intradermal injections which provide continuous, sustained levels of the drug in the patient. Treatment of human patients or other animals will be carried out using a therapeutically effective amount of an anti-pathogenic agent in a physiologically-acceptable carrier. In the context of treating a bacterial infection a "therapeutically effective amount" or "pharmaceutically effective amount" indicates an amount of an antibacterial agent, e.g., as disclosed for this invention, which has a therapeutic effect. This generally refers to the inhibition, to some extent, of the normal cellular functioning of bacterial cells (e.g., Enterococcal cells) causing or contributing to a bacterial infection.

The dose of antibacterial agent which is useful as a treatment is a "therapeutically effective amount." Thus, as used herein, a therapeutically effective amount means an amount of an antibacterial agent which produces the desired therapeutic effect as judged by clinical trial results, standard animal models of infection, or both. This amount can be routinely determined by one skilled in the art and will vary depending upon several factors, such as the particular bacterial strain involved and the particular antibacterial agent used. This amount can further depend on the patient's

height, weight, sex, age, and renal and liver function or other medical history. For these purposes, a therapeutic effect is one which relieves to some extent one or more of the symptoms of the infection and includes curing an infection.

The compositions containing antibacterial agents of virulence factors or genes can be administered for prophylactic or therapeutic treatments, or both. In therapeutic applications, the compositions are administered to a patient already suffering from an infection from bacteria (similarly for infections by other microbes), in an amount sufficient to cure or at least partially arrest the symptoms of the infection. An amount adequate to accomplish this is defined as "therapeutically effective amount." Amounts effective for this use will depend on the severity and course of the infection, previous therapy, the patient's health status and response to the drugs, and the judgment of the treating physician. In prophylactic applications, compositions containing the compounds of the invention are administered to a patient susceptible to, or otherwise at risk of, a particular infection. Such an amount is defined to be a "prophylactically effective amount." In this use, the precise amounts again depend on the patient's state of health, weight, and the like. However, generally, a suitable effective dose will be in the range of 0.1 to 10000 milligrams (mg) per recipient per day, preferably in the range of 10-5000 mg per day. The desired dosage is preferably presented in one, two, three, four, or more subdoses administered at appropriate intervals throughout the day. These subdoses can be administered as unit dosage forms, for example, containing 5 to 1000 mg, preferably 10 to 100 mg of active ingredient per unit dosage form. Preferably, the compounds of the invention will be administered in amounts of between about 2.0 mg/kg to 25 mg/kg of patient body weight, between about one to four times per day.

Suitable carriers and their formulation are described, for example, in Remington's Pharmaceutical Sciences by E.W. Martin. The amount of the anti-pathogenic agent to be administered varies depending upon the manner of administration, the age and body weight of the patient, and with the type of disease and extensiveness of the disease. Generally, amounts will be in the range of those used for other agents used in the treatment of other microbial diseases, although in certain instances lower amounts will be needed because of the increased specificity of the compound. A compound is administered at a dosage that inhibits microbial proliferation.

Combination Therapy

The invention further features a method for treating a patient having a bacterial infection, for example, an Enterococcal infection, by administering to the patient an antibacterial combination therapy that includes a compound identified as modulating expression, at the transcriptional, post-transcriptional, translational, or post-translational levels, of a polynucleotide having at least 50% identity to any of the polynucleotides described herein (or that is substantially identical to a polynucleotide described herein) and one or more antibiotics, including, but not limited to, amikacin, aminoglicosides (e.g., tobramycin), aztreonam, carbenicillin, cephalosporines (e.g., ceftazidime or cefipime), chloramphenicol, gentamicin, levofloxacin, meropenem, piperacillin, tazobactam, tetracycline, vancomycin, and quinolones (e.g., ciprofloxacin), simultaneously or within a period of time (e.g., 14 to 21 days) sufficient to inhibit the growth of the bacteria.

Preferably, the compound and antibiotic are administered within fifteen days of each other, more preferably within five or ten days of each other, and most preferably within twenty-four hours of each other or even simultaneously.

In a related aspect, the invention also features a method for treating a patient having an Enterococcal infection. In this method, a patient is administered (a) a first compound (e.g., a compound that modulates the expression a polynucleotide or polypeptide described herein), and (b) one or more antibiotics (such as amikacin, aminoglicosides (e.g., tobramycin), aztreonam, carbenicillin, cephalosporines (e.g., ceftazidime or cefipime), chloramphenicol, gentamicin, levofloxacin, meropenem, piperacillin, tazobactam, tetracycline, and quinolones (e.g., ciprofloxacin)). If desired, the therapy includes administration of two antibiotics according to standard methods known in the art. Such dual antibiotic combinations most preferably include high-dose tobramycin plus meropenem, meropenem plus ciprofloxacin, or tobramycin (4 μg/ml), or cefipime. Other preferred combinations include piperacillin plus tazobactam, or piperacillin plus ciprofloxacin. The antibiotic and compound combination therapy are preferably administered simultaneously or within a period of time sufficient to inhibit the growth of the bacteria.

In any of the foregoing treatments, the compound and antibiotic included in the combination therapy are preferably administered to the patient as part of a pharmaceutical composition that also includes a pharmaceutically acceptable carrier.

Preferred modes of administration include intramuscular, intravenous, inhalation, and oral administration, or a combination thereof.

The antibacterial combinations of the invention can also be part of a pharmaceutical kit. Preferably, the first compound (e.g., a compound identified as modulating expression, at the transcriptional, post-transcriptional, translational, or post-translational levels, of a polynucleotide or polypeptide having at least 50% identity to any one of the polynucleotide or polypeptide sequences described herein (or that is substantially identical to any one of the polynucleotides or polypeptides described herein)) and the second compound, an antibiotic, are formulated together or separately and in individual dosage amounts.

Combination therapy may be provided wherever antibiotic treatment is performed: at home, the doctor's office, a clinic, a hospital's outpatient department, or a hospital. Treatment generally begins at a hospital so that the doctor can observe the therapy's effects closely and make any adjustments that are needed. The duration of the combination therapy depends on the kind of bacteria being treated, the age and condition of the patient, the stage and type of the patient's bacterial infection, and how the patient's body responds to the treatment. Drug administration may be performed at different intervals (e.g., daily, weekly, or monthly) and the administration of each agent can be determined individually. Combination therapy may be given in on-and-off cycles that include rest periods so that the patient's body has a chance to build healthy new cells and regain its strength.

All publications and patents mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent was specifically and individually indicated to be incorporated by reference.

What is claimed is: